REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

AD-A238 782

2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
4 June 1991	ANNUAL Oct 89 To Mar 91

PARAMETRIC STUDY OF DIFFUSION-ENHANCEMENT NETWORKS FOR SPATIOTEMPORAL GROUPING IN REAL-TIME ARTIFICIAL VISION

5. FUNDING NUMBERS
PE - 61102F
PR 2313
TA A8
GR- AFOSR-FD-90-0001

6. AUTHOR(S)
A.M. Waxman and R.K. Cunningham

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Lincoln Laboratory, MIT
P.O. Box 73
Lexington, MA 02173-9108

8. PERFORMING ORGANIZATION REPORT NUMBER
AFOSR-TR 91 0643
~~AFOSR-TR 91-0642~~

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Dr Genevieve Haddad
Air Force Office of Scientific Research
Directorate of Life Sciences
Building 410
Bolling AFB DC 20332-6448

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 words)

This is the first Annual Technical Summary of the MIT Lincoln Laboratory effort into the parametric study of diffusion-enhancement networks for spatiotemporal grouping in real-time artificial vision. Spatiotemporal grouping phenomena are examined in the context of static and time-varying imagery. Dynamics that exhibit static feature grouping on multiple scales as a function of time and long-range apparent motion between time-varying inputs are developed for a biologically plausible diffusion-enhancement bilayer. The architecture consists of a diffusion and a contrast-enhancement layer coupled by feedforward and feedback connections: input is provided by a separate feature extracting layer. The model is cast as an analog circuit that is realizable in VLSI, the parameters of which are selected to satisfy a psychophysical database on apparent motion.

91-05945

14. SUBJECT TERMS		15. NUMBER OF PAGES
neural networks	long-range apparent motion	46
astrocyte glial networks	spatiotemporal grouping dynamics	16. PRICE CODE
diffusion enhancement	interference suppression	

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
(UNCLASSIFIED)	(UNCLASSIFIED)	(UNCLASSIFIED	SAR

NSN 7540-01-280-5500

Standard Form 298 (Rev. 2-89)
Prescribed by ANSI Std. 239-18

Annual Technical Summary

Parametric Study of Diffusion-Enhancement Networks for Spatiotemporal Grouping in Real-Time Artificial Vision

R. K. Cunningham
A. M. Waxman

1 July 1991

Lincoln Laboratory
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LEXINGTON, MASSACHUSETTS 02173

Prepared for the Department of the Air Force under Contract F19628-90-C-0002.

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
LINCOLN LABORATORY

PARAMETRIC STUDY OF DIFFUSION-ENHANCEMENT NETWORKS FOR SPATIOTEMPORAL GROUPING IN REAL-TIME ARTIFICIAL VISION

ANNUAL TECHNICAL SUMMARY REPORT
TO THE
AIR FORCE OFFICE OF SCIENTIFIC RESEARCH

OCTOBER 1989 — MARCH 1991

R.K. CUNNINGHAM
A.M. WAXMAN
Group 21

4 JUNE 1991

LEXINGTON MASSACHUSETTS

ABSTRACT

This is the first Annual Technical Summary of the MIT Lincoln Laboratory effort into the parametric study of diffusion-enhancement networks for spatiotemporal grouping in real-time artificial vision. Spatiotemporal grouping phenonena are examined in the context of static and time-varying imagery. Dynamics that exhibit static feature grouping on multiple scales as a function of time and long-range apparent motion between time-varying inputs are developed for a biologically plausible diffusion-enhancement bilayer. The architecture consists of a diffusion and a contrast-enhancement layer coupled by feedforward and feedback connections: input is provided by a separate feature extracting layer. The model is cast as an analog circuit that is realizable in VLSI. the parameters of which are selected to satisfy a psychophysical database on apparent motion.

TABLE OF CONTENTS

Abstract iii
List of Illustrations vii

1. INTRODUCTION 1

2. BACKGROUND: PSYCHOPHYSICAL PHENOMENA AND RELATED NEUROBIOLOGY 3

2.1 Psychophysical Results 3
2.2 Biological Considerations 10

3. DIFFUSION ENHANCEMENT BILAYER MODEL 17

3.1 DEB Network Architecture 17
3.2 Relationship to Biological Networks 20
3.3 Alternative Models 21

4. NUMERICAL SIMULATIONS 23

4.1 Overview 23
4.2 Ideal Input Source Function 23
4.3 Diffusion Network in Isolation 25
4.4 CE Network in Isolation 26
4.5 Complete DEB Network 27

5. WORK IN PROGRESS 33

REFERENCES 35

LIST OF ILLUSTRATIONS

Figure No.		Page
1	Static feature grouping.	4
2	Gamma motion.	4
3	Long-range apparent motion.	5
4	Parameters affecting long-range apparent motion.	6
5	Calculated velocity versus spatial separation.	7
6	Split effects.	7
7	Merge effects.	8
8	Terms effect.	9
9	Cat retinal astrocytes.	11
10	Neuroglial-neuronal interactions.	13
11	DEB circuit model.	18
12	TC ideal input.	26
13	TC velocity versus spatial separation.	27
14	Isolated feedback profiles as B decreases.	28
15	SERC ideal input source function.	29
16	Simulated multiscale static grouping.	30
17	Simulated gamma effect.	31
18	Simulated long-range apparent motion.	32

1. INTRODUCTION

This study suggests that diffusion-enhancement interactions play a fundamental role in human preattentive perception. Understanding parallel networks that simulate such interactions is important for understanding neurobiological findings and will suggest new experiments for researchers in that field. Of special interest is the study of spatiotemporal networks because they aid the understanding of the static and dynamic grouping of stimuli that occurs in and across the visual, auditory, and somatosensory systems [1-3]. Such grouping phenomena are also important in artificial vision systems [4]. During the past year, a biologically plausible network model [4-7] has been refined and used to explain some static and dynamic grouping experiments from the psychophysical literature. This diffusion-enhancement bilayer (DEB) represents the interactions between a layer of astrocyte glial cells and a neuronal layer. Long-range communication is achieved via the diffusion of K^+ throughout the electrically coupled glial layer, and percept localization occurs in the contrast-enhancing (CE) neuronal layer. These glial and neuronal layers are coupled by K^+ currents leaking from glial endfeet in close proximity to the neuronal layer.

In order to explain the DEB model, the reader must understand some of the biological and phychophysical results that led to the development of this model, thus a review of some relevant literature is presented before the model itself is described. Numerical simulations performed with the model are also discussed.

The results of examining the leaky diffusion layer in isolation from the CE layer are presented first; some of the abilities of the isolated CE layer are demonstrated next; and work performed with the full network, which reproduces some of the described psychophysical results, is discussed last. The simulation section suggests how the remaining phenomena will be achieved, enumerates some of the problems and successes of the study so far, and discusses work still to be done.

2. BACKGROUND: PSYCHOPHYSICAL PHENOMENA AND RELATED NEUROBIOLOGY

Psychophysical experiments that are being modeled and neurobiology that could support such phenomena are described in this section. Both psychophysics and neurobiology provide data used to set the DEB model parameters.

2.1 Psychophysical Results

2.1.1 Spatiotemporal Grouping Phenomena

The DEB model replicates the effects of both static and dynamic grouping. A striking example of static feature grouping in the visual domain is demonstrated by Marroquin's diagram shown in Figure 1 [8]. Notice how the dots in the diagram appear to group with their neighbors on greater and greater scales over time as one stares at the center of the hexagon.

The simplest dynamic apparent motion effect is known as gamma motion and occurs when a single light is turned on for a brief period of time and then turned off. Although the light is of a fixed spatial extent, the perception is of a light that first expands and then contracts [1], as shown in Figure 2.

New dynamic grouping phenomena emerge when two distinct stimuli interact over time to form the percept of long-range apparent motion. In the human visual system apparent motion can be demonstrated with two lights of fixed spatial extent that are illuminated at distinct times across a fixed spatial separation (Figure 3). With different spatial separations, illumination times, and interstimulus intervals, the light can appear as two separate lights flashing, as one spot that moves smoothly between two real lights, or as one spot that moves smoothly from the first location, jumps, and continues moving smoothly to the second location [1,2]. The most common form of object motion, phi, gives no impression of a particular shape, while the contrasting beta motion appears to be well-defined.

Similar effects can be achieved with tones in the auditory system [3]. Psychologists have examined, in detail, the conditions that will produce these three distinct types of motion and discovered that for fixed-flash durations there is a clear range of onset-to-onset interval (SOA)[1] versus spatial separation that will produce smooth apparent motion (Figure 4). If one shortens the SOA, lights begin jumping rather than smoothly moving between each other, while a still shorter

[1]The onset-to-onset interval is often referred to as the "stimulus onset asynchrony" (SOA) and is defined as the time between the onset of two successive applications of a stimulus. Related is the interstimulus interval (ISI) that is defined as the time between two successive stimuli. Thus ISI + stimulus duration = SOA.

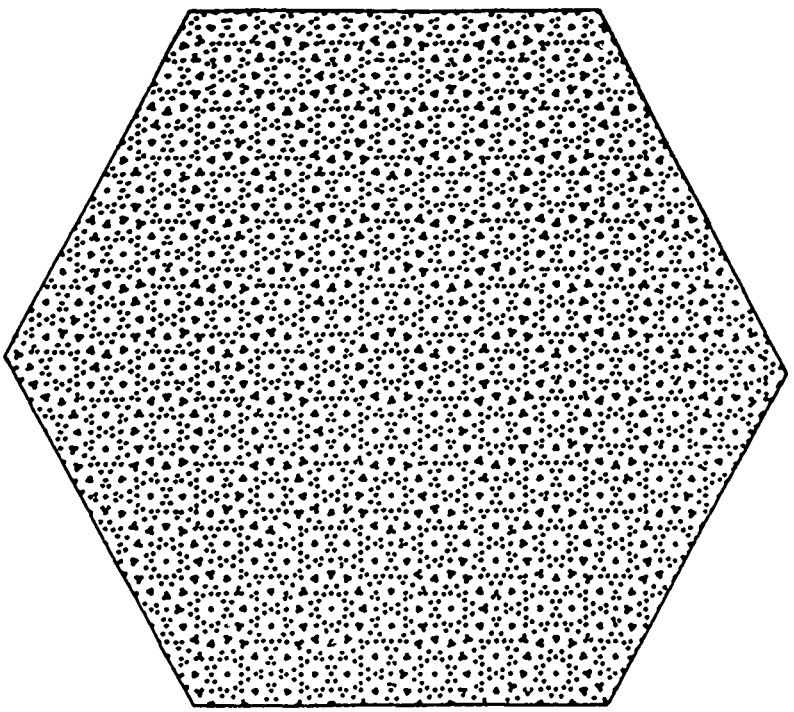


Figure 1. Static feature grouping on multiple scales: evidence for the existence of a grouping process in the visual system [8].

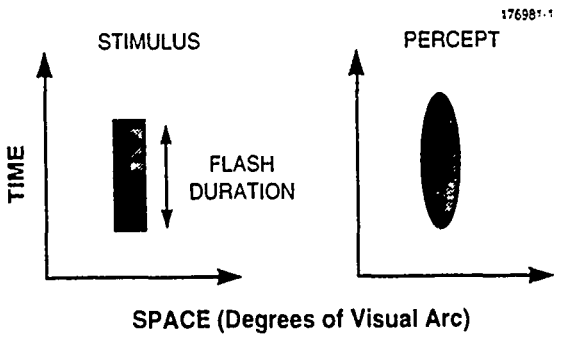


Figure 2. Gamma motion.

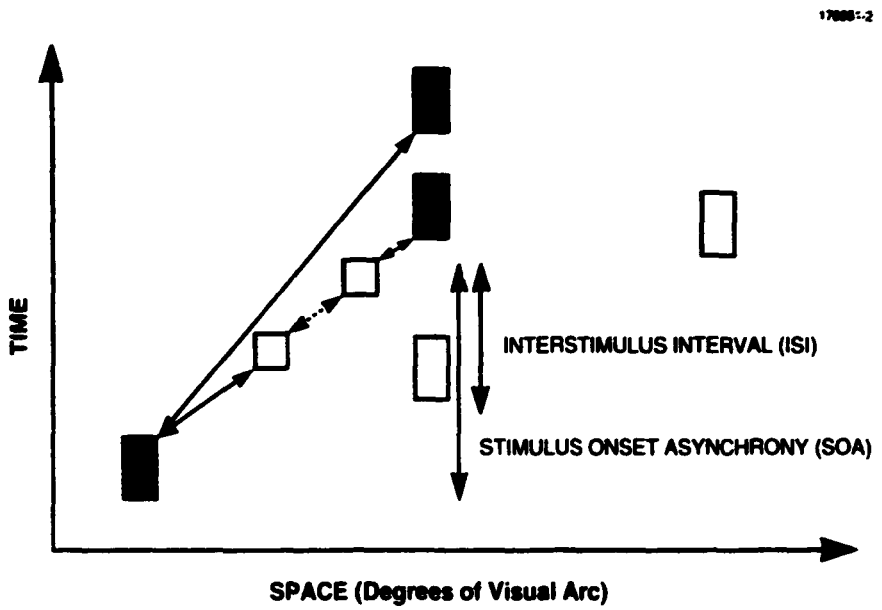


Figure 3. Long-range apparent motion. The filled rectangles represent sources in space-time separations that produce the illusion of long-range apparent motion. The empty rectangles represent sources that are ignited too soon for the spatial distance or are too far away for the given ISI. The center rectangle occurs soon enough to give the appearance of smooth motion from the left-most rectangle to the first shaded square, a short jump to the second shaded square, followed by smooth motion to the destination, while the upper rectangle exhibits pure smooth motion.

SOA causes the lights to appear to flash simultaneously. If one lengthens the SOA beyond an acceptable limit, the lights flash independently of each other. Similar conditions can be created by varying the spatial separation of the two stimuli.

Once this effect was discovered, the next step was to calculate rate of motion. Unfortunately, it is unclear as best to calculate velocity of the illusory figures. Mechanical theory would plot $V = \frac{\text{Distance traveled}}{\text{Time elapsed}}$; in these experiments the distance traveled is clear, but the interpretation of time is not. If ISI is used (the time between when the object was last seen in its initial location and when it showed up in its final location), cases exist for illusory objects with infinite velocity. SOA is the next reasonable choice, but it has been shown that stimulus duration affects apparent motion. Kolers [1] argues that this is as reasonable a plot as any; this report agrees and displays it in Figure 5 for comparison purposes. It is unclear, however, whether the relationship between velocity and spatial separation is linear.

To further understand apparent motion, other psychophysical experiments have been performed; in the visual domain all involve additional stimuli. Split motion effects are demonstrated with three lights of fixed spatial extent: the center light is illuminated and extinguished, followed

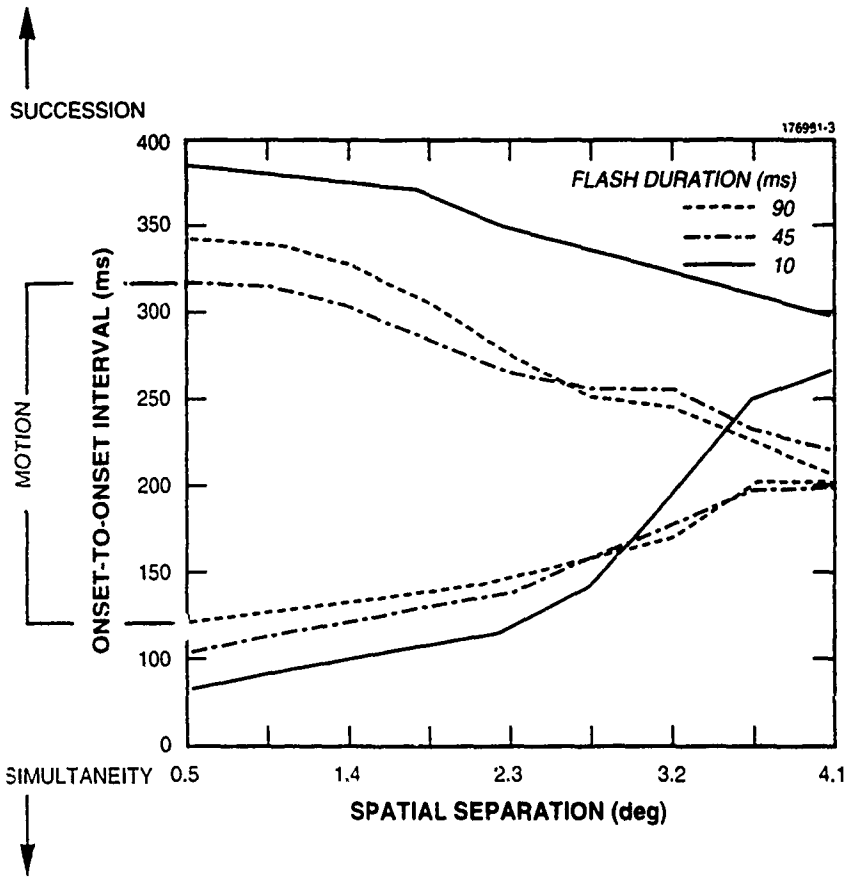


Figure 4. Parameters affecting long-range apparent motion. SOA versus spatial separation.

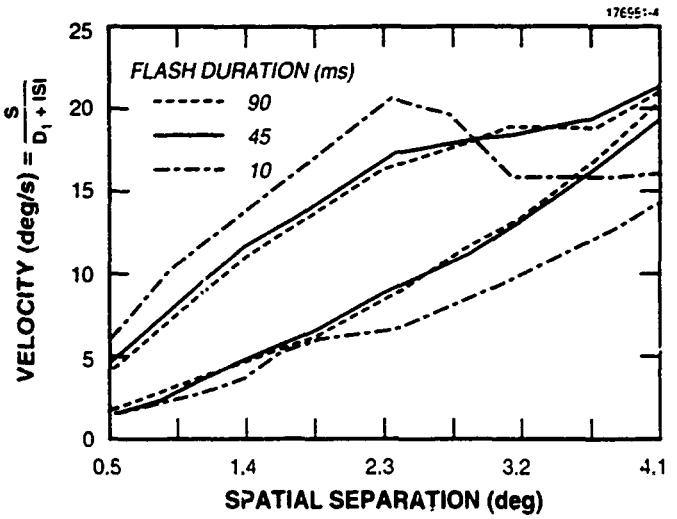


Figure 5. Calculated velocity versus spatial separation [1].

by illuminating and extinguishing the two outer lights. If the two outer lights are equidistant to the center light, the latter appears to split and move to both outer lights. If the two outer lights are at staggered distances, the center light appears to move to the closest light (Figure 6). Also, if the equidistant version of the display is placed in the periphery of the visual field, the center light appears to move towards the light farthest from the fovea.

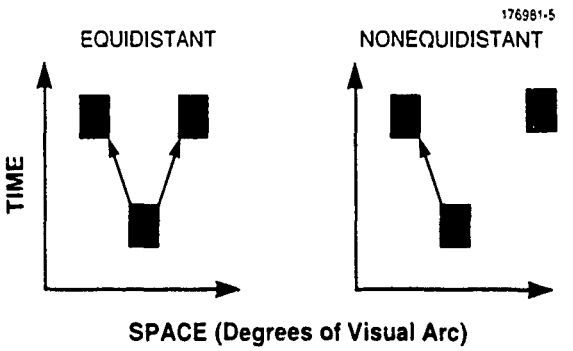


Figure 6. Split effects. Illuminating one point source followed by illuminating two equidistant point sources cause the first to appear to split and move to both of the later sources. If these are not equidistant to the first, movement is only to the closer of the pair.

The opposite of the split effect is the merge effect, in which first the two outer lights, then the center light are illuminated and extinguished. If the two outer lights are equidistant to the center, they both appear to move to the center light and merge with it. If the two outer lights are not equidistant, only the one that is closer appears to move to the center (Figure 7). Once again, if the equidistant display is placed in the periphery of the visual field, the outermost light appears to move toward the center.

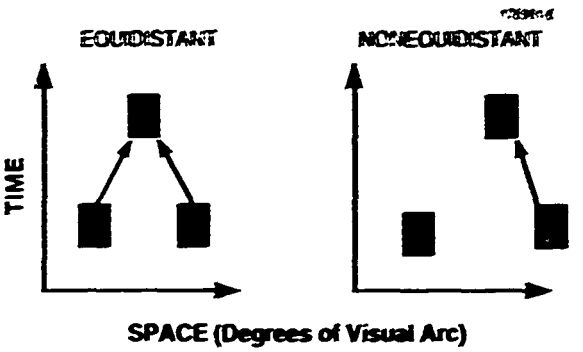


Figure 7. Merge effects. Illuminating two point sources equidistant from a third, later-illuminated point source gives the appearance of the two merging and becoming one. If the original pair is not equidistant from the third point source, only the closer of the two will appear to move.

Another important multielement stimulus studied causes the Ternus effect, which illustrates that more than one distinct motion percept can be achieved from the same display by altering subtle aspects. In the Ternus display, two frames with three aligned lights each are illuminated and extinguished in succession. The frames are aligned so that two of the three lights occupy the same space, and the third appears alternately on the left and right of the central two objects. When the ISI is short, the third light appears to move around the central two objects (element motion), but when the ISI is long the three appear to shift as a coherent group (Figure 8).

Although the cause of these motion effects has been debated for nearly a century, it is known that they do not occur at the retinal level, as evident from a variation on the basic long-range apparent motion experiment. In this variant, known as dichoptic presentations, the first light is presented to the subject's left eye, and the second light is presented to the subject's right eye. Apparent motion is experienced, indicating that spatiotemporal interactions occur at the cortical level. In addition and based on the split and merge experiments performed in the periphery, these effects occur after the visual system has compressed (or down-sampled) the periphery in favor of the fovea, i.e., at the visual cortex. Such sampling and compression occur in at least two places in the

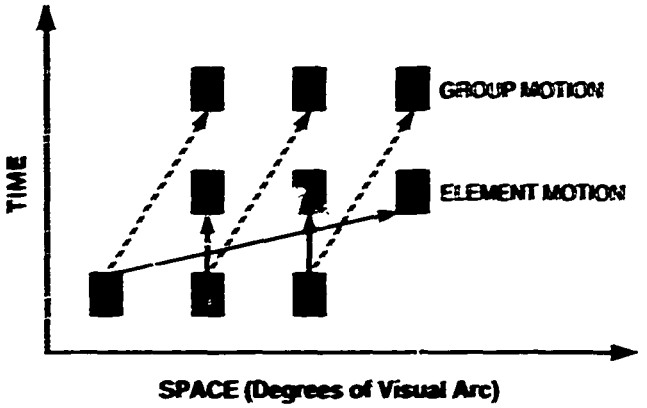


Figure 3. Ternus effect. Illuminating shifting groups of three point sources produces the illusion of one in element motion or all three in group motion. The percept changes with changing ISI.

visual system [9]. First, the retina itself contains a nonuniform population of rods and cones: the densest population occurs in the fovea. Second, the receptive fields of the retinal ganglion cells are markedly smaller in the fovea than in the periphery, hence, the well-known cortical magnification of the fovea and compression of the periphery. Based on neurobiology and psychophysics, it seems evident that the substrate of apparent motion lies in the visual cortex.

This grouping process is not restricted to sight. In the tactile area, if two vibrators agitate the skin with a small ISI, the subject experiences a single vibration between them. As in the visual domain, the effect occurs not in the skin but in the cortex: if the skin between the two vibrators is locally anaesthetized, the effect is still experienced [10]. Similar experiments have been devised for the auditory sense, and similar results were reported. Most bizarre are the intermodal experiments in which apparent motion is perceived between a sound and a light source [3].

2.1.2 Useful Psychophysical Parameters

Psychophysical literature suggests that the structure of the DEB model and long-range apparent motion examples continue to support this diffusion concept and begins to suggest rates at which it should occur. Figure 4 provides spatial and temporal parameters to which the model should conform. The lower curves suggest that communication time varies with distance, while the upper curves suggest that object memory will fade after a time and that this fade occurs faster over greater distances. Both are considered indicative of a leaky diffusion network in which stimulus activity first rises and spatially expands, then falls and contracts: the DEB network was designed to produce these effects. For a fixed (small) spatial separation, there should be a range of SOAs in which motion should be perceived, and outside that range no motion should be realized.

It is difficult to directly relate the model time and space scales to biological time and space, so ratios are examined that allow units to be ignored. One ratio considered is that of the longest to the shortest SOA for which smooth motion occurs at a fixed spatial separation. Figure 4 suggests that this ratio is restricted to ≤ 4 with the ratio decreasing as the spatial separation increases. It is not certain how to compare psychophysical with model-predicted velocities, because there is still no way to interpret the velocities of apparent motion.

2.2 Biological Considerations

Consider that the above psychophysical effects have two salient processes: long-range communication that facilitates interaction of features generated by the inputs (point light sources in the visual domain) and a focusing process that enables objects in apparent motion to have a definitive location. This section discusses the neurobiological elements that are believed most likely to provide the foundations for both these processes. Because it has already been suggested that static and dynamic grouping occurs in the cortex, the search begins there.

The cortex is divided almost evenly between nonneuronal and neuronal cells. It is felt that the largest class of nonneural brain cells, the neuroglial, is the site of the long-range communication process, while neuronal networks provide focusing.

2.2.1 Astrocyte Glial Cells

Once thought of as only providing passive physical support, neuroglial cells now appear to play an "active role in maintaining normal brain physiology" [11]. Concentration is on the astrocyte glial cells because they are known to provide long-range communication between coupled astrocytes (Figure 9). Although to date such coupling has not been directly demonstrated in vivo (Kettenman and Ransom, [2] suggest that this is due to technical difficulties), there is some evidence that it occurs[2] and there is direct evidence for coupling in cultured astrocyte cells. Such communication is not rare. Indeed, Kettenman and others have observed that "mammalian astrocytes in cell culture are highly coupled to one another electrically" and that "qualitative studies have shown that cultured astrocytes form a highly coupled electrical syncytium" [12] that is believed to provide the long-range communication necessary to support the above psychophysical phenomena.

2.2.2 Neuronal Networks

Many cells in the visual cortex are known to derive their input from networks of neurons. In Hubel [9], simple cells (which respond to oriented lines) are postulated to be made up of a hierarchy of lower-order, radially symmetric, center-surround cells. Similarly, complex cells (which respond

[2]Low molecular weight dye passes between adjacent cells [13], and glial networks are postulated to act as potassium spatial buffers [11,12,14].

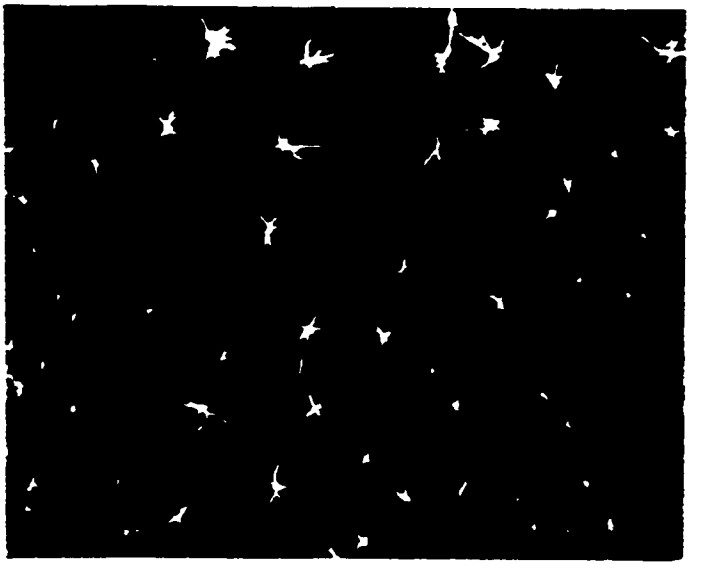


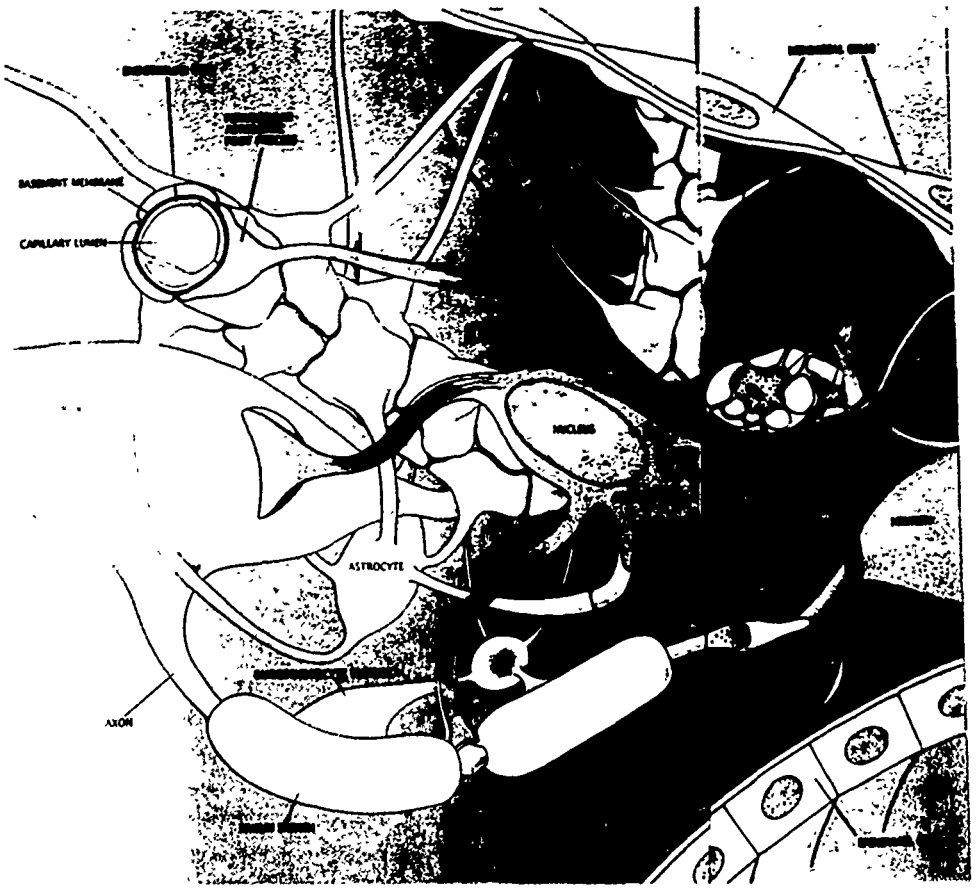
Figure 9. Cat retinal astrocytes. "The radiating fibers enable astrocytes to interact with neurons..." [11].

to oriented lines in a wide receptive field) and end-stopped cells are made up of a network of simple cells. Also, directionally tuned, motion-sensitive cells are postulated to consist of inhibitory and excitatory connections. Similar inhibitory and excitatory interconnections are used in the CE neuronal network where cells are assumed to be self-exciting and latterly inhibiting within the layer.

2.2.3 Interactions

Astrocyte glial cells are known to interact electrically among themselves and with neurons as well; the nature of that interaction is now considered.

Kettenman and Ransom [12] discovered that in cultured astrocyte syncytia, the resistance of the electrical gap junctions between astrocytes is not voltage dependent over much of the membrane potential fluctuation, thus a charge flow model examining intra-astrocyte communication should be independent of the membrane potential of the astrocytes. Futhermore, it is known that "glial cells...have a high potassium concentration and have negligible ionic permeability for ions other than potassium" [15]; therefore, current flow is modeled through glial cells as transfer of potassium ions to or from the cell in a manner obeying Ohm's law. When this process is expanded to encompass current flow in a network of glial cells, the motion of these ions can be approximated with a diffusion equation.



Neuroglial-neuronal interactions. Astrocytes are present near specific ... take up and metabolize specific neurotransmitters [11].

It is further proposed that a glial syncytium provides long-range communication between neurons in a layer via transmission of potassium and other ions. It has been shown that "variations in $[K^+]_{extracellular}$ have profound effects on neuronal excitability, modulating such processes as synaptic transmission and the initiation and propagation of action potentials" [14]. Such $[K^+]$ variations can be realized near the leaky endfeet of astrocytes that are in close proximity to neuronal synapses (see Figure 10). This report is not the first to propose such an interaction. In 1965, Hertz [16] proposed "a mechanism...in which the potassium ions, which have been lost from one nerve cell during its activity, are transported through neuroglia cells to the outer surface of another nerve cell, which is then depolarized and stimulated: that is, a neuronal-neuroglial-neuronal impulse transmission." Hertz continues: "Potassium ions which have been released from an active area are transported through neuroglia cells to the outside of other neurones [sic]: these are in turn stimulated and potassium ions are released, to be transported actively through other neuroglia cells. In this way the spreading depression is propagated across the entire cortex more rapidly than can be explained by a diffusion." The DEB model explicitly uses such interactions to spread and reinforce the charge distribution in a diffusion layer.

2.2.4 Useful Biological Parameters

Odette and Newman [14] note that glial cell endfeet "can contain up to 95% of the total cell conductance." This is important to determine how "leaky" the diffusion process should be.

Kettenmann and Ransom [12] have examined astrocyte coupling by electrically stimulating (via KCl injections) one glial cell and measuring its voltage and that of the neighboring cell. The ratio of these voltages are fit to an exponential, which approximates the steady-state decay in a 1- and 2-D syncytium: $\frac{V_2}{V_1} = \epsilon xp\left(\frac{-d}{L}\right)$, where d is the distance from the injection, and L is the length constant. Kettenman and Ransom [12] measured astrocyte L in vitro to be 80 to 100 μm. L can be used to relate the DEB model to physical size of the biological networks and is related to the ratio of the conductances representing the ratio of G_{gg} to G_g in Section 3.1. Further, L is not related to the 1-D model explained herein as it is believed that the 1-D length constant would have to be significantly greater than the 2-D. Indeed, experiments with restricted 2-D syncytia have L values that are greater than their full 2-D syncytia counterparts [12]. L is expected to be more valuable in the context of the 2-D experiments.

3. DIFFUSION ENHANCEMENT BILAYER MODEL

3.1 DEB Network Architecture

The DEB model consists of two processes that mirror the two salient psychophysical processes mentioned in Section 2.2. i.e., a diffusion layer that facilitates long-range interactions via local connections and a focusing layer that provides the sensation of a localized object traversing a spatial separation. In this model, featural input is presented and preprocessed by the sensor before passing it on to spatiotemporal grouping. In the case of the primate vision system, both center-surround processing and logarithmic spatial mapping occur before grouping begins in the cortex [9]. Following feature extraction, activity is input to the diffusion layer that interacts with a localizing CE layer, which periodically samples the state of the diffusion layer; its output is fed back to the diffusion layer to provide support for new input and facilitate sustained interactions. This report proposes that a motion detection system, such as that of van Santen [17] or Waxman [18], detects the smooth motion of the activity maximum at the output of the CE layer and causes the sensation of motion in the psychophysical experiments. Also, activity prompted by a single input at first grows and eventually dies down, so that after a period of time grouping is no longer possible (Figure 4). This is an effect of the limited time span of featural input from a single feature, the leaky diffusion layer, and the imposition of decay on the feedback from the CE layer.

With this high-level description of the network in mind, a 1-D circuit form of the DEB model is illustrated in Figure 11. Note the two layers — a diffusion layer that permits long-range charge interactions and a CE layer that localizes the charge distribution from the diffusion layer and produces improved SNR via the feedback pathways. Currently, the electrical model is simulated by integrating the governing equations also shown in Figure 11. A separate input layer of feature-sensitive neurons provides activity to the diffusion layer via glial cell endfeet. Determining the profile of this source and comparing it to the biology is a major thrust of our research and is discussed in Section 4.2. Glial endfeet also bidirectionally carry feedforward and feedback activity (i.e., charge or K^+).

The diffusion layer of the DEB model is governed by three coupled differential equations based on Ohm's law; the first represents the spatially coupled diffusion layer:

$$\frac{dQ_g^{(i)}}{dt} = \frac{G_{gg}}{C_g}\left[Q_g^{(i-1)} + Q_g^{(i-1)} - 2Q_g^{(i)}\right] - \frac{G_g}{C_g}Q_g^{(i)} + \frac{G_{ge}}{C_g}\left[\frac{C_g}{C_e}Q_e^{(i)} - Q_g^{(i)}\right] + \frac{G_{gi}}{C_g}\left[\frac{C_g}{C_i}Q_i^{(i)} - Q_g^{(i)}\right]. \quad (1)$$

This equation contains several parameters that can be considered independent of the other coupled equations. Conductivity G_{gg} controls the speed with which charge Q_g is distributed throughout the layer, while G_g controls how rapidly charge leaks into the environment. Together,

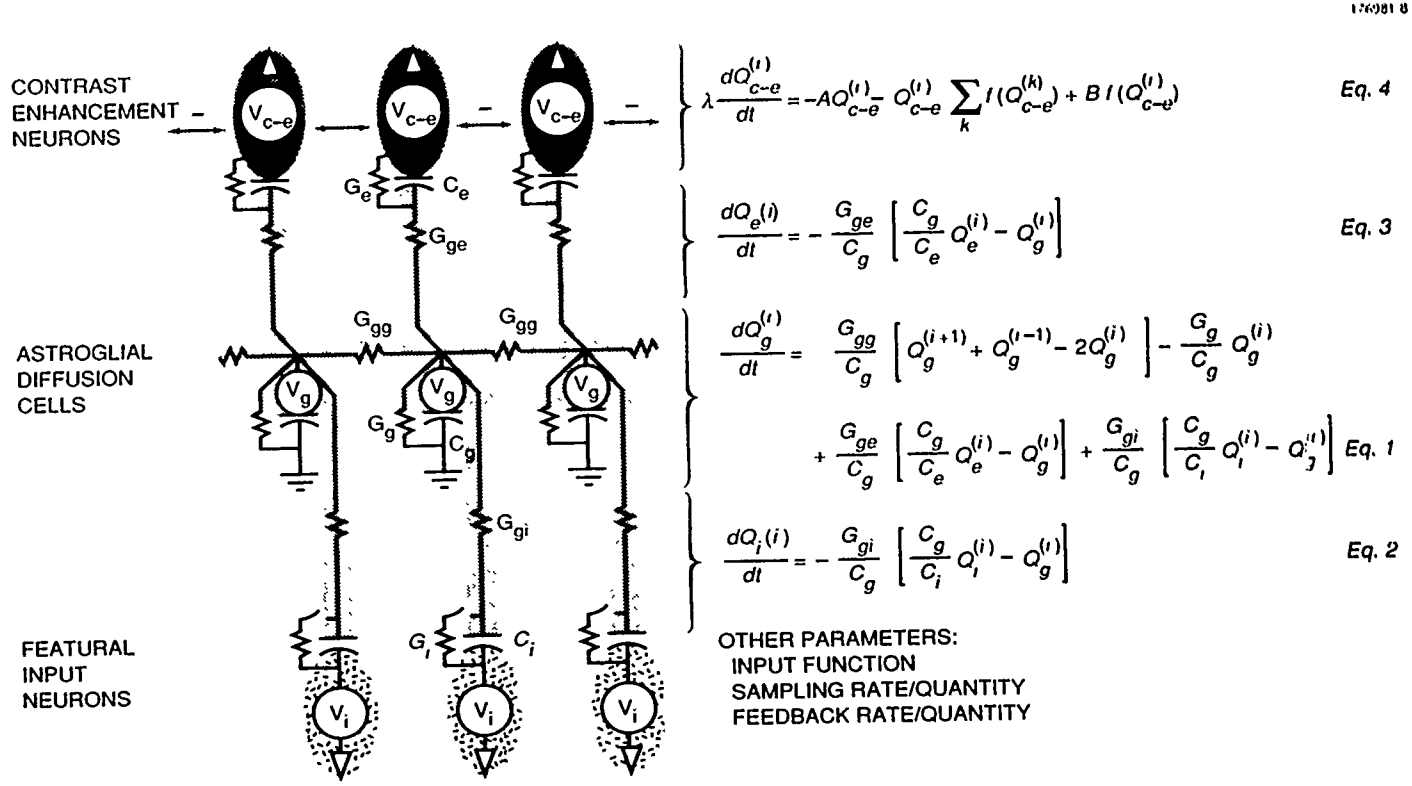


Figure 11. DEB circuit model. Charge is distributed throughout the network via the dynamics in the listed equations. Note the two conceptual layers (and the input) that provide long-range communication (astroglial diffusion cells) and localization (CE neurons).

conductivities G_{gg} and G_g determine the spatial extent over which charge can spread in the diffusion layer. The other two equations govern charge input to the diffusion layer as provided by the featural input neurons [Equation (2)] and feedback from the CE neurons [Equation (3)].

$$\frac{dQ_i(i)}{dt} = -\frac{G_{gi}}{C_g}\left[\frac{C_g}{C_i}Q_i^{(i)} - Q_g^{(i)}\right]. \tag{2}$$

$$\frac{dQ_\epsilon(i)}{dt} = -\frac{G_{g\epsilon}}{C_g}\left[\frac{C_g}{C_\epsilon}Q_\epsilon^{(i)} - Q_g^{(i)}\right]. \tag{3}$$

Conductance G_{gi} controls the rate at which new inputs affect the charge on the diffusion layer. G_{ge} controls the rate at which the CE layer feels the effects of developing charge distributions on the diffusion layer, as well as the rate at which feedback from the CE layer modifies the diffusion layer. The capacitors represented in all three equations store the distribution of charge in the diffusion layer C_g, and at the interfaces to the input C_i and the enhancement layers (C_e).

The charge in the CE endfeet is periodically sampled[3] by the CE neurons, which process activity on a shorter time scale than the diffusion layer. The sampled charge is contrast enhanced via a network originally formulated by Grossberg [19], and the output from this network is fed back to the facing endfeet. The equation governing this system of N neurons is

$$\lambda\frac{dQ_{c-\epsilon}^{(i)}}{dt} = -AQ_{c-\epsilon}^{(i)} - Q_{c-\epsilon}^{(i)}\sum_k f(Q_{c-\epsilon}^{(k)}) - Bf(Q_{c-\epsilon}^{(i)}). \tag{4}$$

where

$$f(Q) = \begin{cases} CQ^2 & Q_1 \quad 0 \le Q \le Q_1 \\ CQ & Q_1 < Q. \end{cases} \tag{5}$$

Equation (4) can be rewritten as a shunting short-term memory model with charge limited to the range $[0, B]$. Depending on the choice of parameters, the rapidly attained equilibrium can either pick the node with maximal charge or contrast-enhance the charge across the layer. The latter properties are of interest here where constant signals are suppressed, noise fluctuations are quenched, and all nodes of nearly maximal activity are enhanced. In any case, the dynamics lead

[3]Periodic sampling in time by the CE neurons can be identified with the refractory period of neurons that are phase-locked in a layer.

to a normalization of activity across the layer, with total equilibrium activity equal to $E = B - \frac{A}{C}$. When in this domain, the nodes for which activities fall below

$$Q(t) = \frac{Q_1}{B - \frac{A}{C}} \tag{6}$$

for a sufficiently large time will be forced to lose all activation, i.e., they will be quenched.

Because feedback reactivates the diffusion layer, even once the original input is off, there is the need to dampen the feedback amplitude over time. Without this step a single light will be sustained in memory forever. This problem is resolved by forcing the parameter B in Equation (4) to decay with time between inputs to the system. When a new input stimulates the visual field, the CE layer is reenergized and B is reset to its maximum value. Between inputs this decay is modeled as

$$B = B_{max} D^t, (D < 1) \tag{7}$$

or

$$B = \Phi\left(B_{max} - t^2\right). \tag{8}$$

where $D \approx< 1$, and $\Phi(x)$ is a threshold linear function equal to x, if $x \geq o$ and 0 otherwise. (To date, most experiments have set $o = 0$.)

3.2 Relationship to Biological Networks

A number of DEB model components have direct correlates in biology. Capacitor C_i represents a neuronal-astroglial interconnect that is locally excited by presented features. A feature-sensitive neuron fires; as it repolarizes K^- is released into the extracellular compartment. (The input source function described in Section 2.1.1 tries to model this charge release.) K^- ion pumps bring K^- onto C_i, which represents a highly permeable endfoot of an astrocyte glial cell (cf. K^--spatial buffering [15]). Here the K^- is freely diffused via ion currents within glia (with membrane capacitance C_g) and forms a network through electrical gap junctions between astrocyte glial cells [12]. The interglial connections are represented by the conductors G_{gg}. A portion of the K^- is diffused out of the cells at endfeet to an upper, CE neuronal layer, which is excited by the increased extracellular K^- concentration. This astroglial-neuronal interconnect is represented by C_e. It is hypothesized that the neuronal layer interacts within itself to contrast-enhance its own activity, further releasing K^- as it fires. It then feeds this contrast-enhanced K^- profile back to the glial layer via the same endfeet, and thereby reinforces the charge distribution in the glial network, particularly near the

charge maximum. The output of the CE layer also provides the basis for the percept of a compact form in smooth motion.

3.3 Alternative Models

One interesting alternative model for long-range apparent motion was proposed by Grossberg and Rudd [20]. Their basic model elements that are responsible for creating continuous motion paths from spatially disparate inputs are very similar to those currently being studied. Essentially, localized inputs (e.g., flashes of light) are assumed to excite a spatially extended Gaussian activation pattern of fixed scale. By combining a preprocessing stage, which detects spatial gradients of brightness with a temporal change detector, their input functions grow and decay over time. When this growth function is used to excite the Gaussian activity pattern, one obtains a fixed-scale Gaussian activity wave with amplitude that grows and then decays. Grossberg and Rudd demonstrate that if spatially separate inputs are flashed at different times, then for an appropriate scale Gaussian the two activity waves will merge into a single activity hump, the maximum of which slides continuously from the position of the initial to that of the final input. They then assume a separate contrast-enhancing process detects this moving maximum.

The DEB model shares the two essential elements of the Grossberg-Rudd model, i.e., a spatially extended response to an input that evolves over time, followed by a CE process that localizes the response; however, where Grossberg and Rudd assume a fixed-scale Gaussian response to an input, the DEB utilizes a diffusion process that responds with increasing scale as a function of time. Both models seek other earlier processes to determine the dynamic nature of the input function responsible for exciting the activity profiles that will interact with one another.

Another alternative model, introduced by Waxman et al. [18], is the short-range motion process, the essential concept of which is the temporal growth and decay of a Gaussian activity wave in response to a transient input. For noninteracting inputs (i.e., a Gaussian wave with scale smaller than the spacing of inputs) this process provides a means to extract directly the speed of moving features. When the features are close to one another the waves do interact, and in so doing they interpolate the trajectory between inputs.

4. NUMERICAL SIMULATIONS

4.1 Overview

To understand the numerical simulations performed over the past year, a brief historical perspective is provided. Prior to this study differential equations (governing a much simpler model) were simulated by reproducing the effects of individual terms. In the case of the diffusion layer, activity was literally spread by using Gaussian convolution. In the case of the CE layer, the feedback point was determined by masking out areas of negative or zero Gaussian curvature and negative mean curvature (where the slope of the activity was zero) and selecting the central point of the surviving, locally maximal plateaus [4-7]. The differential equations were integrated directly, thus removing one possible source of error when simulating the network. In addition, a fully parallel CE network, which will be easier to implement in VLSI than the curvature masking mechanism, has been adopted (and adapted) with the added benefit of providing an adjustable distributed feedback to the current diffusion layer. Moreover, the dynamics of the activity flow along the endfeet that connect the layers are included.

Several tools were developed for understanding the complex dynamics of this network. Among these, the concept of the "ideal input" (discussed in Section 4.2) and how the definition of ideal changed with the understanding of the network and the appreciation of its subtleties. Simulations performed on the diffusion layer in isolation are presented, the CE layer in isolation is discussed, and the interconnected bilayer is considered.

4.2 Ideal Input Source Function

When a feature is detected by the input layer of neurons, they fire for some period of time and release K^+ ions into their environment. Thus, an input is characterized by a charge source that is a function of time. Because the smooth motion of a unique maximum between two input locations is to be modeled (see Figure 3), an ideal input can be generated by searching for that input source function that produces the most rapid motion between the two inputs without producing multiple maxima. This is the essence of the current definition of ideal input, and it remains unchanged throughout the discussion. (In principle, it is preferable to use the actual K^+ release rate from neurons, however, such measurements are not readily available.) Starting from the known profile of a Gaussian distribution of activity on all three layers in proportion to the ratio of the capacitors[4] centered at the first input location, the search began for an ideal input at a displaced location that would satisfy the design constraints. This ideal input is designated as a first-generation ideal input. If this ideal input is used to generate the initial activity profile and reexecute the search at

[4]Experiments not presented here indicated that this was a reasonable approximation to the distribution of activity throughout the network after presentation for a limited period of time.

the second location. a second-generation ideal input is created. This procedure is iterated until the ideal input source function converges.

The ideal input was considered in several different domains in order to better understand each. and its definition was extended as new effects were discovered. The most significant variations are:

1. Tight constraint (TC)
    - Relevant to both the isolated diffusion layer and the entire network
    - Ideal for a range of spatial separations between two inputs
    - Requires a unique activity maximum on the glial enhancement endfeet at every integrated time instant
2. Relaxed constraint (RC)
    - Relevant to entire network
    - Ideal for a range of spatial separations between two inputs
    - Unique maximum fed back from CE layer endfeet
3. Extended relaxed constraint (ERC)
    - Relevant to entire network
    - Ideal for a range of spatial separations
    - Unique maximum fed back from CE layer
    - Maintains final position
4. Simplified extended relaxed constraint ($SERC$)
    - Relevant to entire network
    - Ideal for a single spatial separation
    - Unique maximum fed back from CE layer
    - Maintains final position of activity maximum after it has moved.

The first variation (TC) was designed to examine the conditions that would cause an isolated diffusion network to exhibit long-range apparent motion effects (Figure 3) for a range of spatial separations between inputs. The second variation (RC) was developed from TC when it was considered that the output from the feedback layer should serve as input to a separate motion extraction network. It is relaxed because it is possible to have multiple maxima at the enhancment endfeet of the glial layer between the times when the CE layer samples this layer. In addition, the CE layer quenches low-level noise containing multiple maxima when in the presence of a strong signal.

The third variation (*ERC*) adds the constraint that after the maximum has moved from the first to the second location, it must remain there. This constraint was added when it was noticed that the activity peak, after arriving at the second location, would move back again towards the first feature location. Although this was always present in the *TC* ideal inputs, it was more pronounced when the full network was connected with feedback. It was necessary to choose the minimum charge needed at each time to maintain the maximum at its final position.

Finally, when *ERC* created a difficulty, requirements were simplified (the fourth variation, *SERC*) so that the ideal input would work for a fixed distance. This greatly improved the interactive nature of the software development system: the search time for the *ERC* variation (one generation, 150 time units) is on the order of 4 hr on a Sparcstation 1+, while the *SERC* search time is about 20 min. For reference, 150 time units of network execution time (no search) requires a few seconds to calculate.

4.3 Diffusion Network in Isolation

As noted in Section 1, the reporting period began by examining the isolated leaky diffusion network. Initially it was desirable to verify that network dynamics could replicate phenomena discussed in Section 2.1.1 and to gain an understanding of the effects of the individual network parameters.

Electrical network theory states that the ratio of the capacitors largely determines where charge is stored within the network. It was discovered that by choosing $C_g \approx 10 \times C_i$, $C_g \approx 10 \times C_e$, rapid spatial distribution is possible. When C_g is set an order of magnitude lower, charge is isolated in a small region about the input location. When C_g is set higher, interactions occur so rapidly that features quickly wash out. In addition, simulation time increases drastically since the integration step must then be reduced in the numerical integration routines.

G_g represents the astrocyte glial cell potassium leakage to the extracellular environment, which biology suggests should represent a little more than 5% of cell extracellular conductance and also affects lateral spatial distribution of charge. Independently it was discovered that setting $G_g \approx 0.01 \times$ produces activation profiles with distinct maxima that still equilibrate after a reasonable period of time. Choosing $G_g \approx 0.1 \times G_{gg}$ leads to "washed-out" maxima due to loss to the extracellular compartment, while choosing $G_g \approx 0.001 \times G_{gg}$ causes activity to be trapped entirely in the neural layer.

With parameters set as above, *TC* is used as a tool to determine an ideal input that produces effects similar to the long-range apparent motion effects for spatial separations of 3 to 13 nodes. A first-generation *TC* ideal input is shown in Figure 12. The alternating large/small input is caused by the "greedy algorithm" that administers as much charge as possible at each time unit. Since doubling the number of integration time steps doubles the frequency of this alternation, it is not considered important but rather an artifact of the method used to find the *TC* ideal input. On the other hand, the broader scale scalloping pattern is important and is retained at nearly the same frequency when the number of time steps is doubled. This is due to the constraint that the ideal

input must cause smooth motion for all spatial displacements in the range of 3 to 13 nodes. Each of the smaller displacements arrives at its final destination sooner than the larger displacements, and once it does it no longer influences the calculation, which responds by administering the maximum quantity of charge for the remainder of the displacements, resulting in small spikes of activity. Given this ideal input, the long-range apparent motion effect can be recreated for each separation individually, and the speed at which the single maximum moves from its initial to final location can be examined. The rate of motion is shown in Figure 13. Unfortunately, this is only a first-generation ideal input, so although ISI is known the duration of the initial input must be estimated. Because it is known that the initial profile is similar to an input with that ISI, 150 time units are used.

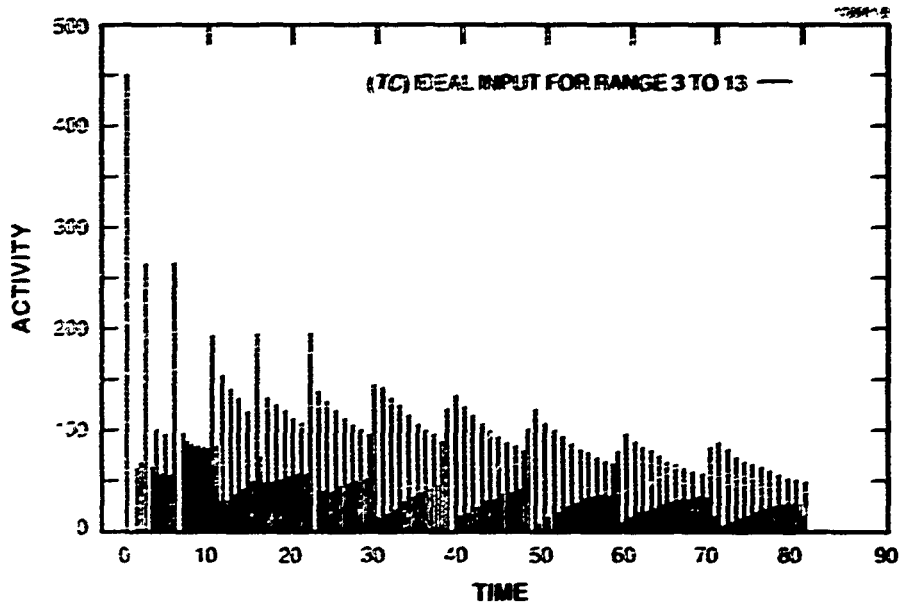


Figure 12: TC ideal input. The scalloped effects are due to each smaller displacement arriving at its final destination.

4.4 CE Network in Isolation

The CE layer provides a localized representation of the diffuse activation. Recall that localized output of this layer also provides the input to motion detectors. In addition, it supports the localized maximum and sustains it in memory, thereby enhancing iterations of inputs to interact on the diffusion layer. After a brief period of time, feedback amplitude decreases and the system forgets past inputs. The process of modifying the amount of activity that is fed back is introduced

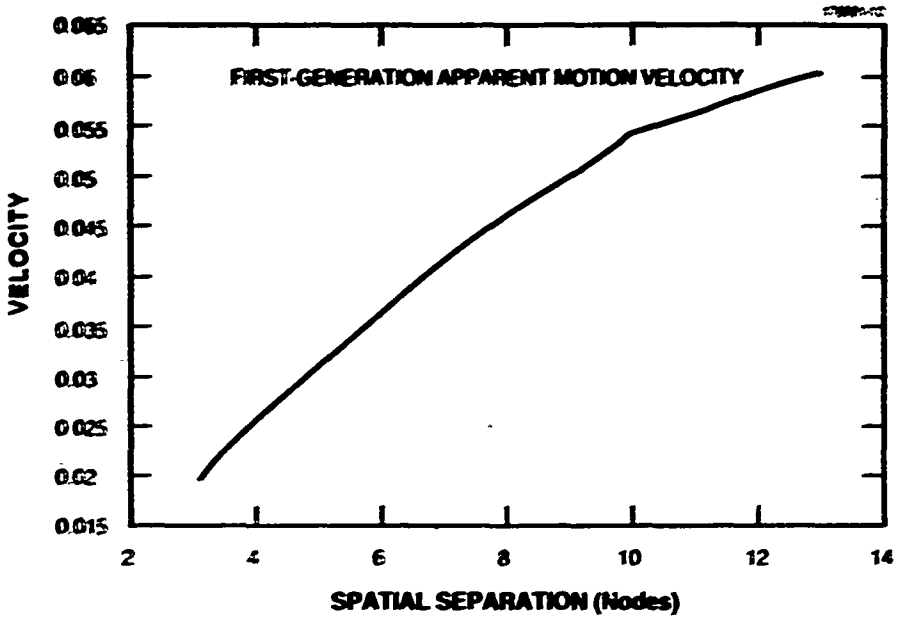


Figure 13. TC: Velocity (nodes/SOA) versus spatial separation. Compare this to curves in Figure 5.

with this study. Before connecting the CE layer to the rest of the system, there were experiments with the effects of a decaying B parameter (see Equations (7) and (8)). The activation profile depicted in Figure 14 comes from an apparent motion experiment in which the initial stimulus occurs at node 50, causes activity growth, and begins to decay as a second stimulus is applied at node 58. The activity profile of the enhancement endfeet (C_e) is sampled when the second stimulus is just beginning to affect the diffusion layer. Note that with a large $B = 5001$, the second input survives the CE layer but with smaller B, e.g., ≥ 1001, the second input is quenched. This property suppresses low-level background noise. Also, note that the extent of the feedback decreases with decreased B. This is important in replicating the effect of an extinguished single light in the gamma motion experiment.

4.5 Complete DEB Network

The next series of examples is the result of working with the entire network. Once the *SERC* ideal input is found, it is used to demonstrate the gamma and the long-range apparent motion effects.

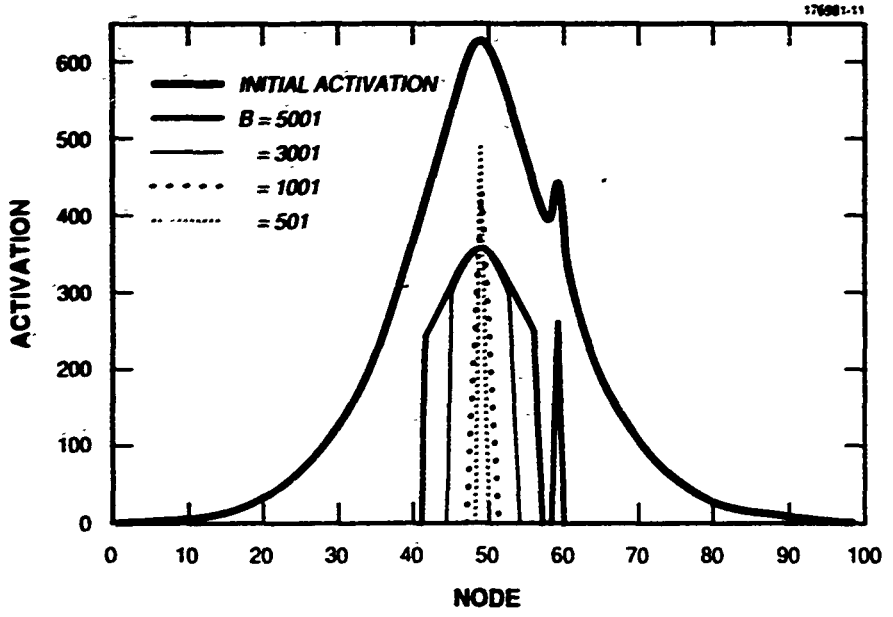


Figure 14. Feedback activation profiles. As B decreases so does the extent of the feedback until the network selects only the maximum node. By causing B to decay between inputs, the network forgets earlier stimulation and reproduces the extinguishing half of the gamma effect.

4.5.1 Ideal Input Source Function

A converged, fourth-generation *SERC* ideal input is calculated for spatial separation of nine nodes using the alternate decay formulation for B [Equation (8)]. The resulting input function is shown in Figure 15 and resembles the firing pattern of a neuron that is depleting its stored K^-. This shape is caused by the combination of two constraints: the greedy algorithm described in Section 4.3 administers as much charge as possible at each time unit, and the evaluation for multiple maxima is enforced only at the output of the CE layer; in this case, every five time units. The second peak in activity occurs in the long-range apparent motion example when the second light source has completed its motion from the first to the second location, and the new criteria (that there must be no backward motion of the activity maximum) engages. This input is used for all the remaining examples.

4.5.2 Static Inputs

Figure 16 represents the response of the CE layer to a static 1-D "image" with structure on multiple scales. In this image there are three point-source inputs at nodes 45, 49, and 58. In the following experiments, sampling by the CE layer occurs every five time units, so only the

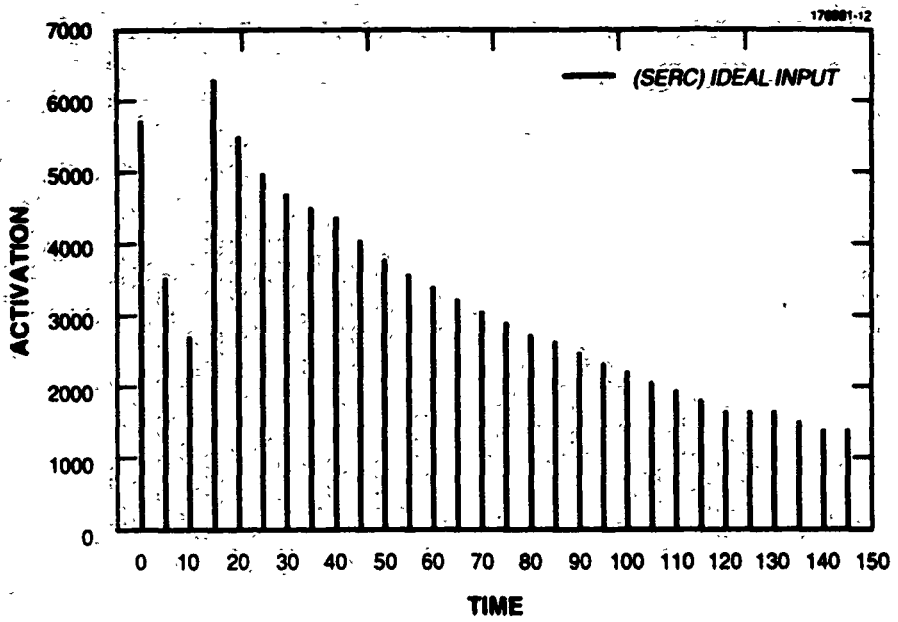


Figure 15. SERC ideal input source function. Converged (fourth generation) for spatial separation of nine nodes between inputs with an ISI of 150 time units.

output from that layer at $t = 5n$. ($n = 1.2.3\ldots$) is illustrated. Excerpts are shown of the CE layer response at increasing time: initially ($t = 5$) the network responds with maxima at each of the three inputs, then ($t = 10$) only two maxima survive as the two closest sources interact and merge on the diffusion layer, and finally ($t = 50$) all three merge, and the network displays only a single maximum on this coarser scale. As $t \rightarrow \infty$ this network generates a maximum that can be used as a focus of attention near the geometric mean of the input locations. Note that if one's eyes followed the absolute maximum value, concentration would first be on the small, then the larger-scale interactions.

4.5.3 Gamma Motion

The next experiment demonstrates the gamma effect (Figure 17). Stimulus is administered to the input endfeet, activity accumulates in the diffusion layer and percolates up to the enhancement endfeet, which activate the CE layer. This layer initially responds with narrow feedback, then spreads the spatial extent of the feedback in response to the spreading of input activity, and finally collapses the spatial extent back to a point-source. The amplitude of the feedback diminishes with time, permitting the system to forget ancient inputs. The result is a blob that expands and contracts with time.

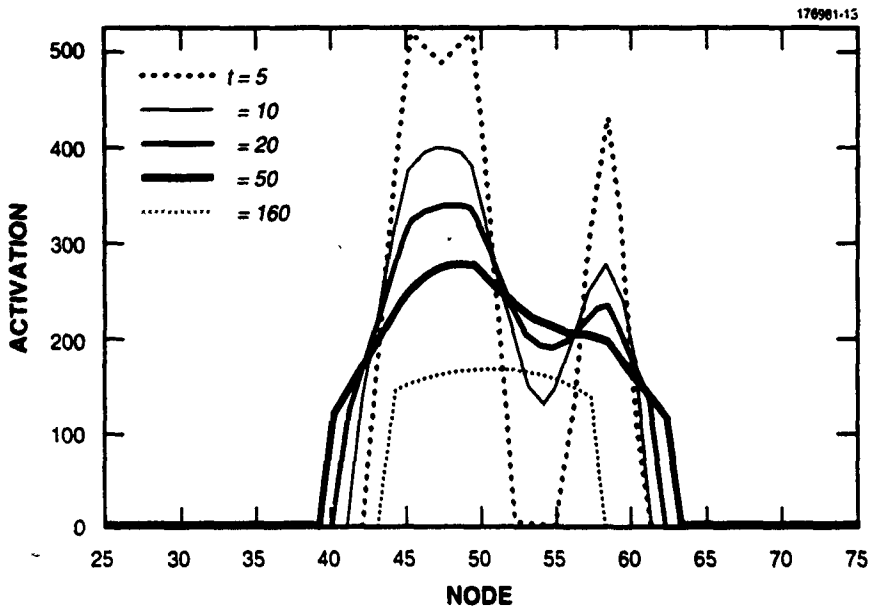


Figure 16. Multiscale grouping. Three inputs at nodes 45, 49, and 58 interact on increasing scale with increasing time. Discrete activation on the CE layer is pictured as continuous curves.

4.5.4 Long-Range Apparent Motion

The dynamic image in the long-range apparent motion experiment described above was examined, concentrating on motion effects. At $t = 0$, a source is activated at node 50, and 150 time units later a second input is activated at node 59. At node 50, the CE layer initially produces a maximum, which moves through node 51, skips 52 to 55, moves through 56, skips 57, and completes its journey at node 58 (Figure 18).

Although input functions that provided smoother motion have been demonstrated, this example is not as smooth as had been hoped. Two possibilities exist: (1) the second input might have occurred too soon after the first, and the system responded with a "jumping maximum" such as the center black rectangle in Figure 3; and (2) the network parameters may not have been correctly set. By increasing $Q1$ in Equation (5), the output of the CE layer can be better localized, which would inhibit motion by producing increased support in the area of the current maximum value, thus causing slower motion that could smoothly move through all intermediate nodes. Later experiments will examine this possibility.

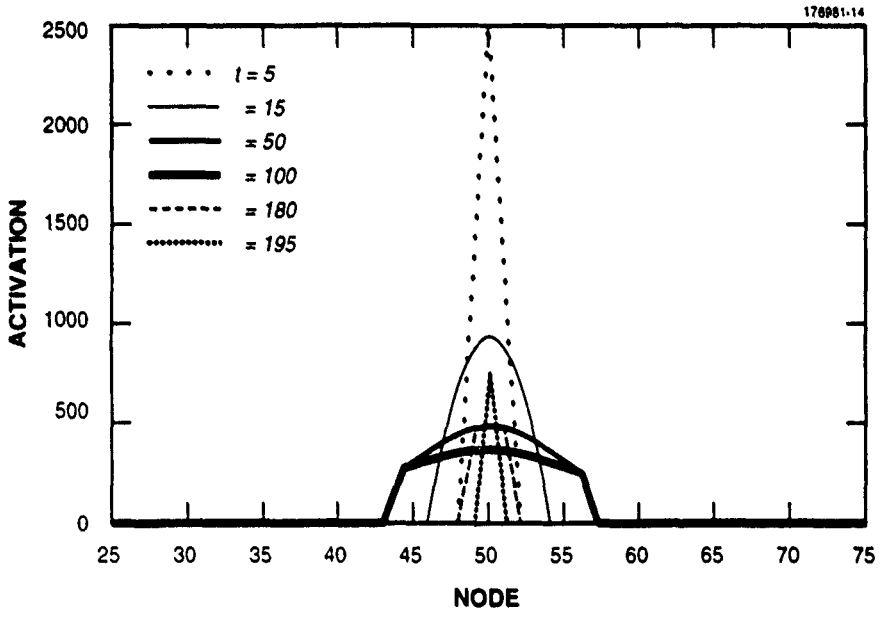


Figure 17. Gamma effect. A single input initially induces a narrow maximum, which expands, stabilizes, and extinguishes after the input is turned off at t = 150. Discrete activation on the CE layer is pictured as continuous curves.

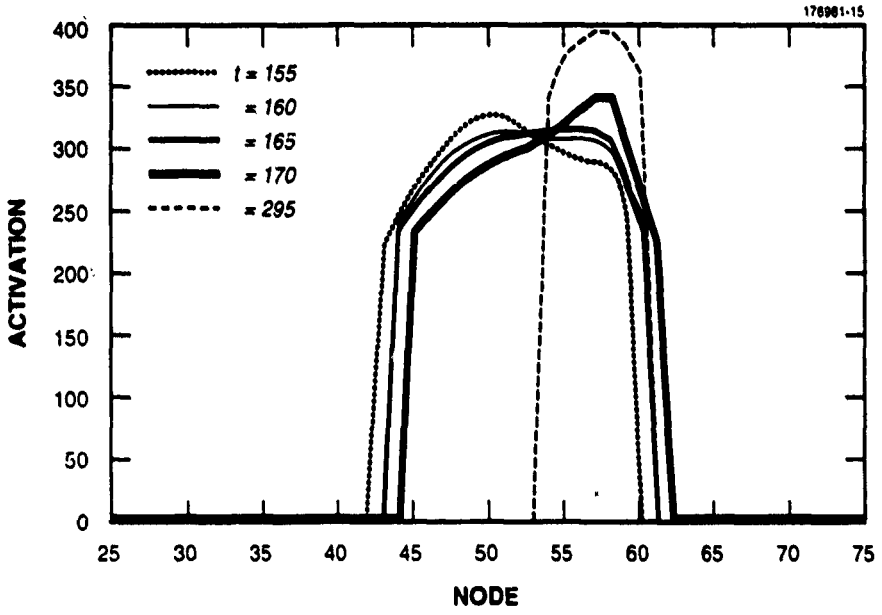


Figure 18. Long-range apparent motion between two inputs. The maximum of activity on the CE layer moves between the initial source located at node 50 and the final source located at node 59, generating the appearance of long-range motion. Discrete activation on the CE layer is pictured as continuous curves. In this case, the maximum of activity exhibits a jump, as in Figure 3.

5. WORK IN PROGRESS

A dynamic model has been introduced for a two-layered network that spatiotemporally groups its inputs on multiple scales as a function of time. Moreover, it generates a long-range apparent motion between spatially separate inputs introduced at different times.

Current research includes a return to the simpler experiments with constant or exponentially decaying input source functions, such as those that were originally performed with the diffusion network in isolation. This approach leads to more rapid understanding of the network and its capabilities while sacrificing the search techniques that lead to an input with known qualities. It had been hoped to find an input that would produce smooth motion over a range of spatial separations for the entire network, and significant time was spent tweaking the system to achieve this end. If this method had been successful, this input function would then have been perturbed to see how it could be modified while still producing the same result. Unfortunately, although initial experiments with the diffusion network in isolation proved successful, a final represention of the input function that would produce smooth motion in the final system was not found. There are a number of possible reasons: the most likely is that a region of space-time outside the smooth-motion regime depicted in Figure 4 was being examined. By using simple input activation profiles such as constant, Gaussian, and exponentially decaying input functions as a guide to the correct domain of parameters, along with the ideal input tool to find an input that will produce smooth motion, it is hoped the DEB model will fully reproduce the psychophysical results listed.

Future work will include the psychophysical phenomena not yet discussed: split, merge, and peripheral effects will all be considered. Once the split and merge effects are working, the peripheral effects should be simple to implement, as these are believed to be due to the topological preprocessing done before the data arrives at the network. To simulate the biological system, the data will be logarithmically mapped before administering activity to the system: this procedure is expected to produce the peripheral psychophysical results. The final effort to be explored is the 1-D Ternus effect: when its results can be reproduced, the DEB model will be extended to two dimensions for oriented inputs.

REFERENCES

1. P.A. Kolers. *Aspects of Motion Perception.* New York: Pergamon Press (1972).
2. M. Wertheimer. Experimentelle studien uber das sehen von bewegung. *Z. fur Psych.* **61**, 161–265 (1912). Translated in part in *Classics in Psychology*, T. Shipley (ed.), New York: Philosophical Library (1961).
3. G.C. Zapparoli and L.L. Reatto. "The apparent movement between visual and acoustic stimulus and the problem of intermodal relations." *Acta Psychologica* **29**. 256–267 (1969).
4. M. Seibert and A.M. Waxman. "Spreading activation layers, visual saccades, and invarient representations for neural pattern recognition systems." *Neural Networks* **2**. 9–27 (1989).
5. A.M. Waxman, J. Wu, and M. Seibert, "Computing visual motion in the short and the long: From receptive fields to neural networks." *Proc. IEEE Workshop on Visual Motion*, Irvine, Calif., 156–164 (1989).
6. A.M. Waxman, M. Seibert, R. Cunningham, and J. Wu. "The neural-analog diffusion-enhancement layer (NADEL) and early visual processing." *Proc. SPIE Conf. on Visual Communications and Image Processing '88*, Cambridge, Mass., 1093–1102 (1988).
7. A.M. Waxman, M. Seibert, R. Cunningham, and J. Wu. "The neural-analog diffusion-enhancement layer and spatiotemporal grouping in early vision." *Advances in Neural Information Processing Systems* **1**. D.S. Touretzky (ed.), San Mateo, Calif.: Morgan Kaufman, (1988).
8. J.L. Marroquin. *Human Visual Perception of Structure*. Master's thesis, Massachusetts Institue of Technology, Cambridge, Mass. (1976).
9. D.H. Hubel. *Eye, Brain and Vision*. Scientific American Library (1988).
10. F.A. Geldard and C.E. Sherrick. "Space time and touch." *Sci. Am.*, 91–96 (July 1986).
11. H.K. Kimelberg and M.D. Norenberg. "Astrocytes." *Sci. Am.*, 66–76 (April 1989).
12. H. Kettenman and B.E. Ransom. "Electrical coupling between astrocytes and between oligodendrocytes studied in mammalian cell cultures." *Glia* **1**. 64–73 (1988).
13. J.A. Conners, L.S. Benardo, and D.A. Prince."Carbon dioxide sensitivity of dye coupling among glia and neurons of the neocortex." *J. Neuroscience* **4**.1324–1330 (1984).
14. L.L. Odette, and E.A. Newman. "Model of potassium dynamics in the central nervous system." *Glia* **1**. 198–210 (1988).
15. S.W. Kuffler, J.G. Nicholls, and R.A. Martin. *From Neuron to Brain*. Sinauer Associates, Inc. (1985).
16. L. Hertz. "Possible role of neuroglia: A potassium-mediated neuronal-neuroglial-neuronal impulse transmission system." *Nature* **4989**. 1091–1094 (1965).

REFERENCES
(Continued)

17. J.P.H. van Santen and G. Sperling. "Elaborated Reichardt detectors." *J. Opt. Soc. America* **A2**, 300–321 (1985).

18. A.M. Waxman, J. Wu, and F. Bergholm. "Convected activation profiles and the measurement of visual motion." *Proc. 1988 IEEE Conf. on Computer Vision and Pattern Recognition*, Ann Arbor, Mich., 717–723 (1988).

19. S. Grossberg. "Contour enhancement, short-term memory, and constancies in reverberating neural networks." *Stud. Appl. Math.* **52**, 217–257 (1973).

20. S. Grossberg and M.E. Rudd, "A neural architecture for visual motion perception: Group and element apparent motion." *Neural Networks* **2**, 421–450 (1989).

21. S. Grossberg. "Adaptive pattern classification and universal recoding II: Feedback, expectation, olfaction, and illusions." *Biol. Cyber.* **23**, 187–202 (1976).

22. S. Grossberg. "Nonlinear neural networks: Principles, mechanisms, and architectures." *Neural Networks* **1**, 17–61 (1988).